

Draft paper demo vegetation indices

R. Willem Vervoort^{*,a}, Dasapta Erwin Irawan^b, Gene Melzack^c

^a*Sydney Institute of Agriculture, The University of Sydney, NSW 2006*

^b*Applied Geology Research Group, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung, Jl. Ganesa No.10, Lb. Siliwangi, Coblong, Kota Bandung, Jawa Barat 40132, Indonesia*

^c*Research Data Team, The University of Sydney Library, The University of Sydney, NSW 2006*

Abstract

Understanding the spatial and temporal variation in primary production can identify variation in ecological response related to climate and landscape. With mobile phone cameras now easily available, this opens an opportunity to use citizen science to record primary production from vegetation images.

This study demonstrates how mobile phone images can be quickly analysed for primary production and can give inside in spatial variation of ecological response.

Participants in an open data workshop took mobile phone imagery of local vegetation and the resulting images were stored in an open data project folder. The analysis calculated the ratio of the green channel and the sum of the red and blue channel from mobile phone images of vegetation.

The results rapid mapping of vegetation vigour and a wide variety across a small area. The process of taking images, uploading, sharing and analysis was rapid using basic open data science tools including the package `rticles`.

Mobile phones can be used by inexperienced operators to quickly map vegetation indices in an area.

Introduction

The evolving nature of science publishing has lead to a strong growth in open science (D. Irawan et al. 2017) and citizen science efforts (Dickinson, Zuckerberg, and Bonter 2010). With this comes a need to create an understanding of open science and the potentials of open science for researchers in different fields. To assist with this a workshop was organised at Institut Teknologi Bandung (ITB) in collaboration with the University of Sydney (USYD) (D. E. Irawan, Vervoort, and Melzack 2018) to demonstrate different aspect of open science and open science publishing.

During the workshop, an example dataset was used based on older research (R. W. Vervoort and Annen 2006). During the workshop, participants were introduced in components, such as defining metadata, creating spatial and netcdfdata, and recording workflows. During the last day of the workshop, new data were generated using vegetation photographs from mobile phones in the field to allow the participants to practice skills learned.

There is considerable interest globally in remotely mapping vegetation due to its strong relationship with ecosystem health and variation in drivers of vegetation productivity (Huete et al. 2011; Rohde, Froend, and Howard 2017). Ratios of colour bands in vegetation photographs have been shown to be highly correlated to gross primary production (Moore et al. 2017), and this therefore creates opportunities for rapid local mapping of vegetation growth and condition.

The objective of this paper is to demonstrate how data can be rapidly collected using mobile phones and the collected data can be easily summarised and a draft paper written using Rmarkdown and the package `rticles`.

*Corresponding Author

Email addresses: willem.vervoort@sydney.edu.au (R. Willem Vervoort), dasaptaerwin@gmail.com (Dasapta Erwin Irawan), gene.melzack@sydney.edu.au (Gene Melzack)

Methods

Data collection

Using mobile phones, data were collected by the participants at several locations on a field site in Bandung, by simply taking photos of the existing vegetation. These photos were downloaded from the phones and collected in a data folder on [the project Google drive storage](#).

Data processing

Metadata were extracted from the photos using the exif program, using the R package `exifr`. This focussed specifically on the latitude and longitude information as this could be used for further mapping. However, other metadata, such as cameratype and time of collection, can also easily be extracted.

Using the package `imager` in R, the image colour bands were extracted from the photos and following Moore et al. (2017) the ratio (GCC) of the green band (G) over the sum of the red (R) and blue (B) bands was calculated:

$$GCC = \frac{G}{(R+B)}$$

Results

Using Rmarkdown, we can run all the analysis in the background and not show this in the paper. However the workflow is still recorded in the code blocks in the “Rmd” file that generates the document. This means the analysis is still repeatable even though this is not directly visible in the article.

As a first result we can show a table of the result with the associated latitudes and longitudes. This can easily be done using `pander` which creates nice tables. This shows that there is quite a range of variability in the data.

Table 1: Table 1 Overview of first 15 collected data points and GCC values for which Latitude and Longitude data are available

| | Latitude | Longitude | GCC |
|-----------|----------|-----------|------|
| 3 | -6.888 | 107.6 | 0.53 |
| 4 | -6.888 | 107.6 | 0.59 |
| 5 | -6.888 | 107.6 | 0.53 |
| 6 | -6.888 | 107.6 | 0.53 |
| 7 | -6.888 | 107.6 | 0.53 |
| 8 | -6.888 | 107.6 | 0.53 |
| 9 | -6.888 | 107.6 | 0.56 |
| 10 | -6.888 | 107.6 | 0.51 |
| 11 | -6.888 | 107.6 | 0.5 |
| 13 | -6.888 | 107.6 | 0.55 |
| 14 | -6.889 | 107.6 | 0.52 |
| 15 | -6.889 | 107.6 | 0.52 |

Some data are not shown, as people did not have their location available on the mobile phone and this means there is no latitude and longitude metadata. This are 33 data points with missing data.

The average **GCC** value for all data is 0.55, while for the data set that does not include the data with missing location data, the average **GCC** value is 0.54. This suggests that there is little effect of omitting the data.

From the plot we can see that collected points are scattered across the area. Overall the variation in GCC is between 0.5 and 0.7. Some of the points are so close together that they are simply overlapping.

Discussion and conclusions

While there are many different ways to publish research. Open Science offers new opportunities to make data and research available to the wider public. Rmarkdown in combination with the package `rticles` allows for efficient recording of workflows and paper writing.

This paper demonstrates that using mobile phone photography it is easy to collect a large amount of data using inexperienced (citizen science) type data collection and to quickly summarize this data.

It has also demonstrated that using existing tools it is fairly simple to produce a draft paper that includes the overall workflow within the document.

References

- Dickinson, Janis L., Benjamin Zuckerberg, and David N. Bonter. 2010. "Citizen Science as an Ecological Research Tool: Challenges and Benefits." *Annual Review of Ecology, Evolution, and Systematics* 41 (1): 149–72. doi:[10.1146/annurev-ecolsys-102209-144636](https://doi.org/10.1146/annurev-ecolsys-102209-144636).
- Huete, Alfredo, Kamel Didan, Willem van Leeuwen, Tomoaki Miura, and Ed Glenn. 2011. "MODIS Vegetation Indices." In *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*, edited by Bhaskar Ramachandran, Christopher O. Justice, and Michael J. Abrams, 579–602. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-6749-7_26.
- Irawan, Dasapta E, Rutger W Vervoort, and Gene Melzack. 2018. "Open Data Workshop Sseac Usyd - Itb." OSF.io. <https://OSF.IO/S76GU/>.
- Irawan, Dasapta, Cut Novianti Rachmi, Mochammad Multazam, Hendy Irawan, Juneman Abraham, kustiat, KeuKeu Kaniawati Rosada, et al. 2017. "A Review on the Implementation of Open Science in Indonesia." OSF.io. <https://osf.io/preprints/inarxiv/7r8jn/>.
- Moore, C. E., J. Beringer, B. Evans, L. B. Hutley, and N. J. Tapper. 2017. "Tree-grass Phenology Information Improves Light Use Efficiency Modelling of Gross Primary Productivity for an Australian\hack\newline Tropical Savanna." *Biogeosciences* 14 (1): 111–29. doi:[10.5194/bg-14-111-2017](https://doi.org/10.5194/bg-14-111-2017).
- Rohde, Melissa M., Ray Froend, and Jeanette Howard. 2017. "A Global Synthesis of Managing Groundwater Dependent Ecosystems Under Sustainable Groundwater Policy." *Groundwater* 55 (3): 293–301. doi:[10.1111/gwat.12511](https://doi.org/10.1111/gwat.12511).
- Vervoort, R. W., and Y. L. Annen. 2006. "Palaeochannels in Northern New South Wales: Inversion of Electromagnetic Induction Data to Infer Hydrologically Relevant Stratigraphy." *Soil Research* 44 (1): 35–45. <https://doi.org/10.1071/SR05037>.

List of Figures

| | | |
|---|--|---|
| 1 | Plot of the collected data by latitude and longitude | 5 |
|---|--|---|

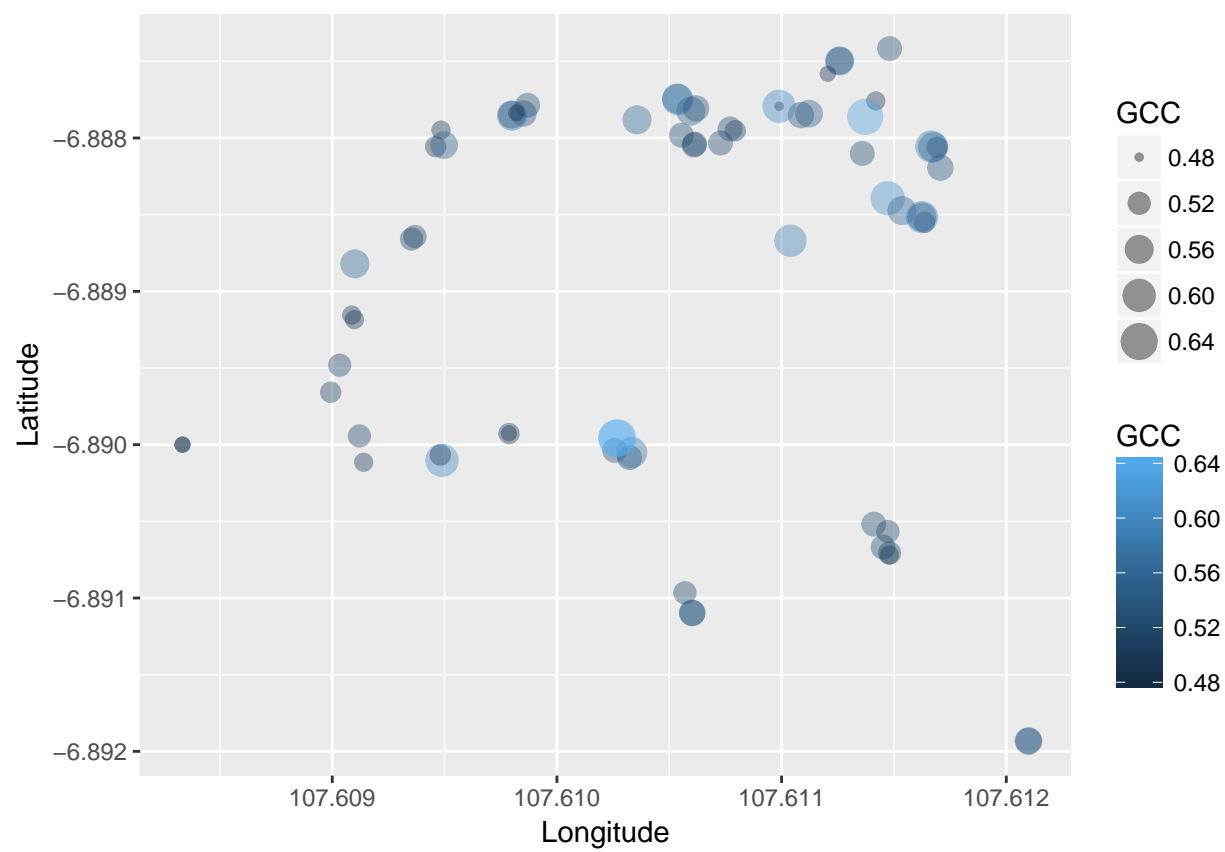


Figure 1: Plot of the collected data by latitude and longitude